Systems/Circuits

Relating Divisive Normalization to Neuronal Response Variability

Ruben Coen-Cagli[1,2] and Selina S. Solomon[2]

[1]Department of Systems and Computational Biology, and [2]Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, Bronx, New York 10461

Cortical responses to repeated presentations of a sensory stimulus are variable. This variability is sensitive to several stimulus dimensions, suggesting that it may carry useful information beyond the average firing rate. Many experimental manipulations that affect response variability are also known to engage divisive normalization, a widespread operation that describes neuronal activity as the ratio of a numerator (representing the excitatory stimulus drive) and denominator (the normalization signal). Although it has been suggested that normalization affects response variability, we lack a quantitative framework to determine the relation between the two. Here we extend the standard normalization model, by treating the numerator and the normalization signal as variable quantities. The resulting model predicts a general stabilizing effect of normalization on neuronal responses, and allows us to infer the single-trial normalization strength, a quantity that cannot be measured directly. We test the model on neuronal responses to stimuli of varying contrast, recorded in primary visual cortex of male macaques. We find that neurons that are more strongly normalized fire more reliably, and response variability and pairwise noise correlations are reduced during trials in which normalization is inferred to be strong. Our results thus suggest a novel functional role for normalization, namely, modulating response variability. Our framework could enable a direct quantification of the impact of single-trial normalization strength on the accuracy of perceptual judgments, and can be readily applied to other sensory and nonsensory factors.

Key words: divisive normalization; modeling; neuronal variability; single-trial inference; visual cortex

Significance Statement

Divisive normalization is a widespread neural operation across sensory and nonsensory brain areas, which describes neuronal responses as the ratio between the excitatory drive to the neuron and a normalization signal. Normalization plays a key role in several important computations, including adjusting the neuron's dynamic range, reducing redundancy, and facilitating probabilistic inference. However, the relation between normalization and neuronal response variability (a fundamental aspect of neural coding) remains unclear. Here we develop a new model and test it on primary visual cortex responses. We show that normalization has a stabilizing effect on neuronal activity, beyond the known suppression of firing rate. This modulation of variability suggests a new functional role for normalization in neural coding and perception.

Introduction

Variability is a widespread feature of neuronal activity in sensory cortex: spike counts to repeated presentations of the same sensory input are variable (Tolhurst et al., 1983; Arieli et al., 1996). Traditionally, studies of neural coding have focused on the mean response averaged over repetitions (i.e., the firing rate) to extract the signal (differences in activity between conditions) from the noise (changes across repeats of the same condition). In this view, variability limits neural coding. However, just like firing rate, variations of neuronal activity across trials also contain meaningful information. For instance, stimulus onset reduces variability in many cortical areas (Churchland et al., 2010). In early visual cortex, variability is modulated along several stimulus dimensions: for instance, contrast (Orbán et al., 2016), size (Snyder et al., 2014), orientation (Goris et al., 2014), and motion direction (Ponce-Alvarez et al., 2013). Nonsensory signals, such as attention (Cohen and Maunsell, 2009; Mitchell et al., 2009; Harris and Thiele, 2011; Ecker et al., 2014; Rabinowitz et al., 2015) and locomotion (Bennett et al., 2013; Dadarlat and Stryker, 2017), also

Received Jan. 15, 2019; revised June 13, 2019; accepted June 18, 2019.

Author contributions: R.C.-C. designed research; R.C.-C. and S.S.S. performed research; R.C.-C. analyzed data; R.C.-C. wrote the first draft of the paper; R.C.-C. and S.S.S. edited the paper; R.C.-C. wrote the paper.

S.S.S. was supported by Fight for Sight. The data were collected in the Kohn laboratory, with support from National Institutes of Health Grants EY026240 and EY016774. We thank Adam Kohn for helpful discussions and for comments on the manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Ruben Coen-Cagli at ruben.coen-cagli@einstein.yu.edu.

https://doi.org/10.1523/JNEUROSCI.0126-19.2019

modulate neuronal variability. Many of these effects persist even after accounting for firing rate differences (Churchland et al., 2010), suggesting variability may be a genuine coding dimension of neuronal activity.

There is evidence that modulation of variability may depend on divisive normalization (Albrecht and Geisler, 1991; Heeger, 1992), a canonical operation evident across multiple brain areas (Carandini and Heeger, 2011). First, experimental factors that affect response variability, such as stimulus contrast, size, and attentional modulation, are also thought to control the strength of normalization (Heeger, 1992; Schwartz and Simoncelli, 2001; Reynolds and Heeger, 2009; Coen-Cagli et al., 2015). Second, numerical simulations demonstrate that normalization can strongly modulate one aspect of response variability, namely, how this variability is shared between pairs of neurons (i.e., noise correlations) (Tripp, 2012; Verhoef and Maunsell, 2017). Furthermore, although the mechanisms of normalization are not fully known, inhibitory stabilization of network activity (Rubin et al., 2015; Hennequin et al., 2018) recapitulates many effects described by normalization, further pointing to a link between changes in variability and in effective normalization strength.

However, we lack a descriptive model of the relation between normalization and variability. Existing normalization models describe only the firing rate, and thus cannot assess the effect of normalization on variability. On the other hand, successful descriptive models of neuronal variability (Paninski, 2004; Pillow et al., 2008; Goris et al., 2014; Charles et al., 2018) typically ignore normalization. We therefore develop a model that explicitly parametrizes this relation and allows us to estimate the relevant parameters from data. The normalization equation describes the firing rate of a neuron as the ratio between the stimulus drive to its receptive field, and the stimulus drive to an ensemble of other neurons (termed normalization signal). We treat the numerator and the normalization signal as two random variables, and derive analytical estimators for the across-trial mean and variance of their ratio, as well as an efficient method to infer the single-trial value of the normalization signal. The model reproduces Poisson-like behavior and rate-dependent Fano factors consistent with known cortical data (Tolhurst et al., 1983; Goris et al., 2014). We then apply the model to macaque primary visual cortex (V1) responses to gratings of varying contrast because contrast modulation of firing rate is well described by the standard normalization model (Heeger, 1992; Tolhurst and Heeger, 1997). Our model captures accurately both the mean and variance of the spike count as a function of contrast, including a systematic reduction of the Fano factor at high contrast (Orbán et al., 2016), which other popular descriptive models fail to capture. The model predicts that increasing normalization strength often reduces variability. Consistent with this prediction, we find that V1 neurons that are more strongly normalized respond more reliably, and V1 activity is less variable and less correlated during trials in which the normalization signal is inferred to be stronger.

Our results thus indicate a general relation between normalization and variability. The model can be extended to quantify similar relations for other stimuli and nonsensory factors, beyond contrast in V1. Furthermore, it can be used to better understand perception and behavior (e.g., to quantify the impact of normalization on perception) by comparing perceptual choices with the single-trial normalization strength.

Materials and Methods

The model

The Divisive Normalization model and the Ratio-of-Gaussians (RoG) model. In the generic formulation of the normalization model, the average spike count of a neuron is described by the following:

$$R = \frac{N}{\epsilon^2 + D} \tag{1.1}$$

where the numerator N usually represents the stimulus drive to the neuron; D is the normalization signal, typically conceptualized as the summed activity of a large group of neurons termed normalization pool; N and D are assumed non-negative; and the constant ϵ^2 prevents division by zero.

We extended the normalization model to account for response variability by treating both N and D as Gaussian-distributed random variables (see Fig. 1*a*), and adding Gaussian noise to the ratio (to capture spontaneous activity fluctuations), leading to the RoG model as follows:

$$\begin{cases} R_t = \dfrac{N_t}{D_t} + \eta_t \\ (N_t, D_t) \sim \mathcal{N}\left((\mu_N, \mu_D), \begin{bmatrix} \sigma_N^2 & \rho\sigma_N\sigma_D \\ \rho\sigma_N\sigma_D & \sigma_D^2 \end{bmatrix} \right) \\ \eta_t \sim \mathcal{N}(\mu_\eta, \sigma_\eta^2) \end{cases} \tag{1.2}$$

where the subscript t indexes a single trial. We have implicitly absorbed ϵ^2 in μ_D, and we assume $\mu_D > 0$; $\mu_N > 0$ in the standard normalization model; this constraint makes the probability of negative values of R low. Therefore, when fitting the model (detailed below) to non-negative count data, parameter values that assign high probability to negative R would lead to a bad fit (i.e., low likelihood of the data), and will be automatically discouraged. Nonetheless, negative values of R have nonzero probability under the RoG, whereas the firing rate is strictly not negative. In our simulations with realistic parameters, we discarded the rare instances with negative single-trial count.

Approximated moments of the RoG distribution. Analytical expressions for the RoG distribution (Hinkley, 1969; Pham-Gia et al., 2006) and other ratio distributions (Springer, 1979; Pham-Gia and Turkkan, 2005) have been obtained, but they involve complicated expressions that make it difficult to perform optimization and inference, and to find the moments of the distribution. A well-known example is the ratio of two standard normal variables (i.e., the Cauchy distribution) whose mean and variance are not even defined. This is important because we are interested in characterizing the effects of normalization on response variability. For our purposes, a useful result is that, if D has negligible probability mass at or <0, the RoG is closely approximated by a Gaussian (Marsaglia, 1965, 2006; Pham-Gia et al., 2006). In our context, this assumption on D is well motivated, as D is usually thought of as the summed activity of a large normalization pool of neurons, and so unlikely to be zero or negative. Under these conditions, the RoG output is Gaussian distributed and therefore fully characterized by its mean and variance, for which we can derive approximations using a Taylor expansion around (μ_N, μ_D) as follows:

$$\begin{aligned} R &= \eta_t + \frac{\mu_N}{\mu_D} + \left.\frac{\partial R}{\partial N}\right|_{(\mu_N,\mu_D)} (N - \mu_N) + \left.\frac{\partial R}{\partial D}\right|_{(\mu_N,\mu_D)} (D - \mu_D) + \ \ldots \\ &\simeq \eta_t + \frac{\mu_N}{\mu_D} + \frac{(N-\mu_N)}{\mu_D} - \frac{\mu_N(D-\mu_D)}{\mu_D^2} \end{aligned} \tag{1.3}$$

(neglecting higher-order terms) and taking the expectations, thus obtaining the following expressions:

$$\begin{cases} \mu_R \simeq \dfrac{\mu_N}{\mu_D} + \mu_\eta \\ \sigma_R^2 \simeq \dfrac{\mu_N^2}{\mu_D^2}\left(\dfrac{\sigma_N^2}{\mu_N^2} + \dfrac{\sigma_D^2}{\mu_D^2} - 2\dfrac{\rho\sigma_N\sigma_D}{\mu_N\mu_D} \right) + \sigma_\eta^2 \end{cases} \tag{1.4}$$

We validated the approximations in simulations and found that both remained unbiased (within 0.3%) over several orders of magnitude (see

Fig. 1*b*). σ_R^2 is mainly determined by σ_N^2 and μ_D^2 (i.e., the first term in the parentheses), whereas σ_D^2 has a much smaller effect because it is small relative to the mean (per our assumption that D is positive). The effect of σ_D^2, however, is important, as it allows for single-trial inference of D, as we explain below.

Analytical relation between spike-count variance, stimulus drive, and normalization strength. Equation 1.4 allows us to address directly the question of how the average normalization strength (μ_D) affects response variability. First, we address the case with no additive noise ($\mu_\eta = 0$, $\sigma_\eta^2 = 0$). We quantify variability by the Fano factor as follows:

$$F_R \doteq \frac{\sigma_R^2}{\mu_R} = \frac{\mu_N}{\mu_D}\left(\frac{\sigma_N^2}{\mu_N^2} + \frac{\sigma_D^2}{\mu_D^2} - 2\frac{\rho\sigma_N\sigma_D}{\mu_N\mu_D}\right) \tag{1.5}$$

and we study its derivative with respect to μ_D:

$$\frac{\partial F_R}{\partial \mu_D} = -\frac{\sigma_N^2}{\mu_N\mu_D^2} - 3\frac{\sigma_D^2\mu_N}{\mu_D^4} + 4\frac{\rho\sigma_N\sigma_D}{\mu_D^3} \tag{1.6}$$

If N and D are uncorrelated ($\rho = 0$), the derivative is always negative, implying that the Fano factor decreases with the strength of normalization. When $\rho > 0$, the derivative can change sign. To find the amount of correlation required for this to happen, we set the derivative to zero:

$$\frac{\partial F_R}{\partial \mu_D} = 0 \Leftrightarrow \mu_D = \left(2\rho \pm \sqrt{4\rho^2 - 3}\right)\frac{\sigma_D\mu_N}{\sigma_N} \tag{1.7}$$

This shows that a solution exists only if $\rho^2 \geq 3/4$; therefore, an increase in normalization strength can increase the Fano factor only for very large correlation levels between N and D. Also from Equation 1.5, regardless of the value of ρ, the dependence of the Fano on the numerator mean is not monotonic in general because it is the sum of a term proportional to μ_N and one proportional to μ_N^{-1}.

Next, we consider the case with $\sigma_\eta^2 \neq 0$. For simplicity, we set $\mu_N = 0$, and also $\rho = 0$ because it has a very limited influence per our analysis above. In this case, the Fano factor is as follows:

$$F_R = \frac{1}{\mu_D\mu_N}\sigma_N^2 + \frac{\mu_N}{\mu_D^3}\sigma_D^2 + \frac{\mu_D}{\mu_N}\sigma_\eta^2 \tag{1.8}$$

This expression shows that the relative sizes of the additive noise and of the variances of D and N determine whether the Fano factor increases or decreases as a function of either μ_D or μ_N.

Reproducing Poisson-like variability. A well-established property of cortical response variability *in vivo* is that it is Poisson-like (i.e., the variability increases approximately monotonically with the mean spike count) (Tolhurst et al., 1983). For the RoG distribution, changes in mean and variance (Eq. 1.4) are controlled by changes in the parameters of Equation 1.2, encompassing a broader range of variance-mean relationships. To further constrain the RoG parameters, we reasoned that both N and D are described as signals of neural origin, and thus we assumed Poisson-like variability for each of those terms separately, following a commonly used formulation (Tolhurst et al., 1983; Carandini et al., 1997):

$$\sigma_{N,D}^2 = \alpha_{N,D}\mu_{N,D}^{\beta_{N,D}} \tag{1.9}$$

For $\beta = 1$, the parameter α determines the Fano factor, whereas different values of β allow for nonconstant Fano. We used this relation separately for the numerator and denominator of the RoG model, and this does not imply necessarily that also the variability of R follows the same relation. However, we found empirically that for a wide range of parameter settings the variability of R is Poisson-like (see Fig. 1*c*). Also, if we substitute Equation 1.9 into Equation 1.5, there is no longer a unique, monotonic relation between the Fano factor and μ_D; in particular, for large enough values of β_D, the Fano factor can increase (rather than decrease) as μ_D increases.

Inference of the single-trial normalization signal

An advantage of the Gaussian approximation for the RoG is that it greatly simplifies inference. Specifically, we are interested in inferring the normalization strength D in a single trial (which cannot be measured directly), given the measured spike count R. To show how this is achieved, here we assume that all the parameters (the means and covariance of N and D) are known: we explain in the next section the model-fitting procedure to estimate those parameters from data.

We are interested in the posterior distribution of D given R, which can be obtained by Bayes rule as follows:

$$p(D|R) = \frac{p(R|D)\,p(D)}{p(R)} \tag{1.10}$$

Here p(D) is the Gaussian prior assumed by the RoG model on D. Note that $p(R)$ does not depend on the value of D (only on its mean and variance, which are fixed); therefore, we can ignore the denominator. We first consider the case $\rho = 0$ and no additive noise ($\eta = 0$). Using the fact that $p(R|D) = \mathcal{N}\left(\frac{\mu_N}{D}, \frac{\sigma_N^2}{D^2}\right)$, with straightforward calculations, we obtain the following:

$$\log p(D|R) = \log(|D|) - \frac{1}{2}\left\{\frac{(R^2\sigma_D^2 + \sigma_N^2)D^2 - 2(R\mu_N\sigma_D^2 + \mu_D\sigma_N^2)D}{\sigma_N^2\sigma_D^2}\right\} + \text{constant} \tag{1.11}$$

To compute the maximum a posteriori (MAP) estimate of D (i.e. our estimate of the normalization strength, denoted by D^{MAP}), we search for the zeros of the derivative of eq. (1.11) with respect to D (we assume $D > 0$):

$$\frac{\partial \log p(D|R)}{\partial D} = \frac{1}{D} - \frac{(R^2\sigma_D^2 + \sigma_N^2)D - (R\mu_N\sigma_D^2 + \mu_D\sigma_N^2)}{\sigma_N^2\sigma_D^2}$$

which equals zero only for one value of $D > 0$:

$$D^{\text{MAP}} = \frac{1}{2}\left(\frac{R\mu_N\sigma_D^2 + \mu_D\sigma_N^2}{R^2\sigma_D^2 + \sigma_N^2}\right) + \sqrt{\frac{1}{4}\left(\frac{R\mu_N\sigma_D^2 + \mu_D\sigma_N^2}{R^2\sigma_D^2 + \sigma_N^2}\right)^2 + \frac{\sigma_N^2\sigma_D^2}{R^2\sigma_D^2 + \sigma_N^2}} \tag{1.12}$$

Lastly, we can use the Laplace approximation, and approximate log p ($D|R$) by a Gaussian distribution centered at D^{MAP} with variance $\sigma_{D|R,\text{Laplace}}^2 = \left(\frac{1}{(D^{MAP})^2} + \frac{R^2\sigma_D^2 + \sigma_N^2}{\sigma_N^2\sigma_D^2}\right)^{-1}$ which measures the uncertainty around the estimate D^{MAP}. The asymptotic behavior of eq. (1.12) provides some insight into the conditions under which the inference of D is informative. In the limit of small prior variance σ_N^2 for the numerator, the MAP estimate tends to μ_N/R. In other words, in this condition, trial-by-trial fluctuations in spike count are entirely explained by fluctuations in the normalization strength, and therefore the normalization strength can be estimated accurately. Conversely, in the limit of small prior variance σ_D^2 for the denominator, the MAP estimate converges to the prior mean μ_D, and therefore the estimate of the normalization strength is entirely uninformative (i.e. it is, in each trial, identical to the average across trials). Note also that in trials with $R = 0$, the RoG requires that also $R = 0$ and D cannot be estimated accurately; for this reason, in our analysis of single-trial normalization strength in the experimental data we excluded trials with no spikes (3807/12800 trials).

The case with $\rho \neq 0$ is similar except that $p(R|D)$, and therefore $p(D|R)$, also depend on ρ:

$$\log p(D|R) = -\frac{1}{2}(\Lambda_{NN}R^2 + 2\Lambda_{ND}R + \Lambda_{DD})D^2 + [\Lambda_{NN}\mu_N R + \Lambda_{ND}(\mu_D R + \mu_N) + \Lambda_{DD}\mu_D]D + \log(|D|) + \text{const.} \tag{1.13}$$

where we denote with Λ the inverse of the covariance matrix between N and D. As we explain below, in our application to V1 data we found that including ρ did not improve the model's predictive power due to overfitting, therefore we only used eq. (1.12) for the single-trial inference. However using eq. (1.13) would be just as easy in other data that warranted the use of ρ.

Unfortunately the case with $\eta \neq 0$ does not have such a simple solution. The MAP estimate of D could instead be found numerically by gradient descent, but we did not pursue this further. Nonetheless, we found in simulations that the estimator of D (eq. (1.12)) remains approximately unbiased (within 0.01%) for moderate levels of additive noise, as long as the mean of such additive noise was correctly accounted for (i.e. removing the additive noise mean from the stimulus-evoked spike counts). For this reason, in our analysis of experimental data, we measured the spontaneous rate and removed it from the evoked responses before inferring the single-trial value of D. We also verified that our results were noisier, though not qualitatively different, when we did not remove the spontaneous rate.

Model fitting

Maximum likelihood parameter estimation. We optimized the RoG parameters by maximum likelihood. In the general setting, the RoG model has five free parameters (μ_N, μ_D, σ_N, σ_D, ρ) per stimulus condition, as they can be stimulus-dependent, plus (μ_η, σ_η). The Poisson-like assumption of Equation 1.9 reduces the stimulus-dependent parameters to three (μ_N, μ_D, ρ), plus (α, β), which are stimulus-independent. In our application to the contrast response function data, we found that ρ leads to overfitting; therefore, we set it to zero, and we further parametrized the stimulus dependence of the means (details below). If we denote generically the free parameters by Θ, the negative log-likelihood of a dataset of measured responses $\{R_t(s)\}$ (where t denotes the trial and s the stimulus value) is as follows:

$$
\begin{aligned}
\log p(\{R_t(s)\};\Theta) &= \text{constant} \\
&+ \sum_{s=1}^{S}\sum_{t=1}^{T}\left\{\frac{1}{2}\log[\sigma_R^2(s,\Theta)] + \frac{[R_t(s)-\mu_R(s,\Theta)]^2}{2\sigma_R^2(s,\Theta)}\right\} \\
&\simeq T\sum_{s=1}^{S}\left\{\frac{1}{2}\log[\sigma_R^2(s,\Theta)] + \frac{\langle[R_t(s)-\mu_R(s,\Theta)]^2\rangle_t}{2\sigma_R^2(s,\Theta)}\right\} \\
&= \frac{T}{2}\sum_{s}\left\{\log[\sigma_R^2(s,\Theta)] + \frac{\widehat{\sigma_R^2}(s) + [\widehat{\mu_R}(s)-\mu_R(s,\Theta)]^2}{\sigma_R^2(s,\Theta)}\right\} \qquad (1.14)
\end{aligned}
$$

where in the third line we have dropped the constant; T is the number of trials (identical for each stimulus) and S the number of stimuli; $\langle \cdot \rangle_t$ denotes the expectation over trials; ($\widehat{\mu_R}(s)$,$\widehat{\sigma_R^2}(s)$) denotes the empirical mean and variance across trials, per stimulus condition; and ($\mu_R(s,\Theta)$, $\sigma_R^2(s,\Theta)$) are the model predictions. To find the best fit parameter values, we minimize the negative of the log-likelihood iteratively using MATLAB's "fmincon."

By inspecting Equation 1.4 for the variance, it can be noted that the additive noise acts as a regularizer in Equation 1.14 (i.e., preventing the model variance to become too small). One important difference between fitting the RoG and fitting the standard normalization model is that in the RoG both the predicted mean and the variance of the spike count are stimulus-dependent and so both appear in the objective function, whereas in the standard model only the mean is stimulus-dependent and the optimization reduces to a nonlinear least-squares problem.

Cross-validated goodness of fit. We quantified the ability of the model to capture the data on a test dataset distinct from the training set used for optimization. The training set contains all stimulus conditions and all but one repetition; this held-out repetition, for all stimuli, forms the test set. We performed leave-one-out cross validation (i.e., we averaged the test-set results across all possible splits of the data between training and test set). The measure of goodness of fit (also known as predictive power) is the average log-likelihood computed on the test data. To map this value on an interpretable scale, we normalized it between a null model and an oracle model as is done often (Stocker and Simoncelli, 2006; Goris et al., 2014). The null model predicts that the response mean and variance are stimulus-independent and equal to the mean and variance across all stimuli and repetitions in the training set. The oracle model predicts that the mean and variance in each stimulus condition are identical to the means and variances estimated from the training set. The normalized, cross-validated goodness of fit approaches zero when the model is as bad as the null model, and approaches 1 when it is as good as the oracle.

Parametrization of the contrast response function. To fit the RoG model, we need to specify how $\mu_R(s,\Theta)$, $\sigma_R^2(s,\Theta)$ depend on the stimulus and the parameters, which in turn requires parametrizing the stimulus dependence of the moments of N and D. Here we apply the RoG to contrast response function data. Therefore, we start from the standard normalization formulation, which provides excellent fits to the average spike count as a function of contrast (Heeger, 1992) as follows:

$$R^{DN}(s) = R_0 + \frac{R_{\max}s^2}{\epsilon^2 + s^2} \qquad (1.15)$$

Here, s denotes contrast of the visual stimulus, and (R_0, $R_{\max}$, ϵ) are free parameters. In our experimental data, multiple orientations of the visual stimulus were presented to the neurons, but we fit the model only to responses to a fixed orientation (the one that elicits maximal neuronal response, details below). Therefore, contrast s is the only relevant stimulus dimension. Based on Equation 1.15, we parametrize the RoG model such that the predicted mean response (Eq. 1.4) is identical to the prediction of the normalization model as follows:

$$\begin{cases} \mu_N = R_{\max}s^2 \\ \mu_D = \epsilon^2 + s^2 \\ \mu_\eta = R_0 \end{cases} \qquad (1.16)$$

and we have additional free parameters for the noise variance σ_η^2, as well as (ρ, α_N, β_N, α_D, β_D) from Equations 1.2 and 1.9. We set R_0 to the average spontaneous activity, and we constrain the parameter range as follows: $\epsilon \in [1,100]$; $\alpha_{N,D} \in [.1,20]$; $\beta_{N,D} \in [1,2]$; $\rho \in [-1,1]$; R_{max} between 0.5 and 2 times the maximum response across all stimuli; and σ_η^2 between 0.1 and 10 times the measured variance of the spontaneous activity. In practice, we found that treating σ_η^2 as a free parameter slightly improved predictive power while changing little from its measured values. On the contrary, treating ρ as a free parameter produced strong overfitting (average goodness of fit on the training set 0.82, CI [0.8 0.85]; on the test set 0.65, CI [0.64 0.68]). Therefore, we report below the results with $\rho = 0$ rather than as a free parameter.

Alternative models

The model comparison of Figure 2 uses the popular modulated Poisson model (Goris et al., 2014), which has been shown to capture well the deviations of cortical variability from Poisson statistics (i.e., from constant Fano factor of 1). In brief, the modulated Poisson model assumes that spike counts evoked by a fixed stimulus are Poisson distributed, with a rate determined by the product of a constant stimulus function (the tuning curve) and a fluctuating gain factor (termed modulator). Across-trial fluctuations of the modulator induce super-Poisson variability, leading to Fano factors that increase with the mean spike count. In Figure 2, we used the model of Goris et al. (2014) for the spike-count distribution; but instead of the measured tuning curves, we used the divisive normalization model of Equation 1.15 as the tuning curve. We then fit jointly all parameters (those of the normalization model for the mean tuning curve, plus the variance of the gain factor). This formulation of the modulated Poisson model has fewer free parameters than the RoG. For this reason, we compare the two models by cross-validation, which automatically accounts for the difference in model complexity.

Experimental data and statistical analysis

Animal preparation and data collection. Data were collected from 4 adult male monkeys (*Macaca fascicularis*). Animal preparation and general methods were described previously (Jia et al., 2013). In brief, anesthesia was induced with ketamine (10 mg per kg of body weight) and maintained during surgery with isoflurane (1.0%–2.5% in 95% O_2). During recordings, anesthesia was maintained by sufentanil citrate (6–18 μg per

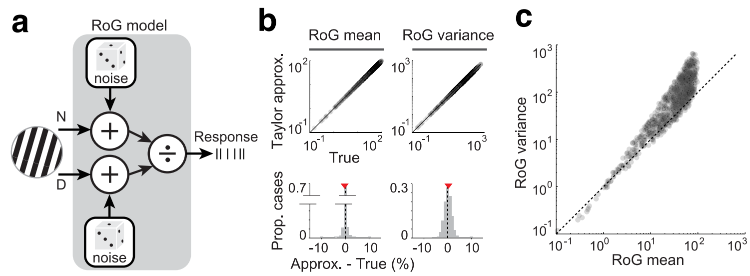

Figure 1. RoG model. ***a***, The RoG model describes neuronal responses as the ratio of two stimulus-driven signals: the numerator *N* and denominator, or normalization signal, *D*. Both *N* and *D* are corrupted by additive Gaussian noise with Poisson-like variability (Eq. 1.9). ***b***, Comparison of the approximation we derived for the RoG mean and variance (Eq. 1.4) and the true value (estimated from 1 million simulated trials), across 1000 experiments. Each experiment uses a different set of model parameters (i.e., corresponding to different neurons and stimuli), and each trial is a random draw from the corresponding distribution. Red triangles represent average percentage difference between true and approximated mean (left) or variance (right). ***c***, RoG variance versus mean across 1000 simulated experiments. Model parameters for the simulations in ***b*** and ***c*** are drawn uniformly in the following intervals: $\mu_N \in [0,100]$; $\mu_D \in [0.5, 1.5]$; $\alpha_N = 1$; $\alpha_D = 0.01$; $\beta_N = \beta_D \in [1,1.5]$; $\rho \in [0,0.5]$; $\mu_\eta = 0$; $\sigma_\eta^2 = 0.1\mu_N/\mu_D$.

kg per hour, adjusted as needed for each animal). Vecuronium bromide (0.15 mg per kg per hour) was used to minimize eye movements. All procedures were approved by the Albert Einstein College of Medicine and followed the guidelines in the United States Public Health Service Guide for the Care and Use of Laboratory Animals.

We recorded neuronal activity using arrays of 10 × 10 microelectrodes (400 μm spacing, 1 mm length) inserted in the opercular region of V1. Waveform segments that exceeded a threshold (a multiple of the root mean square noise on each channel) were digitized (30 kHz) and sorted offline. For all analyses, we included signals from well-isolated single units as well as small multiunit clusters, and refer to both as neurons.

Visual stimuli and presentation. We displayed stimuli on a calibrated CRT monitor (1024 × 768 pixels, 100 Hz frame rate, ~40 cd m^{-2} mean luminance) placed 110 cm from the animal, using custom software. We first measured the spatial receptive field of each neuron, using small gratings (0.5° in diameter, four orientations, 250 ms presentation) presented at a range of positions. Stimuli for the main experiment were then centered on the aggregate receptive field of the population.

To measure contrast response functions, we presented drifting gratings (1° diameter) at 4 orientations and 5 contrast levels (6.25%, 12%, 25%, 50%, 100%). We also added a 0 contrast condition to measure spontaneous activity. Stimuli were presented for 500 ms (3 animals) or 800 ms (1 animal), separated by a blank screen of equal duration, repeated 20–25 times, and interleaved in pseudo-random order. For each neuron, we analyzed contrast response functions only at the best orientation tested.

Characterization of neuronal responses and inclusion criteria. We quantified neuronal activity on each trial by the spike count measured in a window starting 50 ms following stimulus onset (to account for response latency) and ending 50 ms after stimulus offset. Due to the small stimulus size (1°) relative to the typical spread of receptive field centers on the array (~2.5° at the recorded eccentricity), many neurons were not driven by the stimuli. Therefore, we included in the analysis only neurons that were visually responsive ($n = 172$), defined by their average firing rate to the best stimulus being larger than the spontaneous firing rate plus 2 SDs.

As a control, we also considered shorter count windows, starting 50 ms following stimulus onset and ending after 75, 125, 250, or 350 ms. All our results remained qualitatively similar, but the number of neurons and trials that could be analyzed was progressively reduced for shorter windows, for two reasons: (1) with lower spike counts, the Gaussian assumption of the RoG model results in a poorer approximation and therefore worse fit quality; as a result, fewer neurons could be included in the model-based analyses, where we included only neurons with goodness of fit > 0.5; and (2) the proportion of trials with zero spikes increased, thus reducing the number of cases that could be analyzed because the single-trial denominator cannot be estimated accurately in trials with zero spikes, as explained above.

Statistical analysis. CIs reported throughout the text are 95% intervals estimated on 10,000 bootstrap samples. When averaging Fano factor across neurons or conditions, we computed the geometric mean. Significance and p values were obtained with parametric paired t test when comparing goodness of fit for different models and in Table 1, and with a permutation test for the Spearman correlation coefficient.

Code accessibility
Code for model simulations, fitting and inference is available without restrictions on GitHub (https://github.com/rubencoencagli/RoG).

Results
We developed a descriptive model to quantify how normalization affects response variability. To this aim, we extended the standard normalization model, in which the spike count of a neuron is described by the ratio between the stimulus drive to the neuron (the numerator, N) and the normalization signal (the denominator, D). We treated both N and D as random variables (Fig. 1*a*). Specifically, we assumed Gaussian noise in both N and D, and thus obtained a ratio of two Gaussian variables (RoG). In this model, the mean and variance of the spike count are not independent: they both follow from the statistics of N and D. There is no simple closed-form expression for the RoG distribution; but if D has negligible probability mass at or <0 (i.e., positive mean and relatively small variance), the RoG is closely approximated by a Gaussian (Marsaglia, 1965, 2006; Pham-Gia et al., 2006). Under this condition, we derived analytical expressions to approximate the mean and variance of the RoG (Eq. 1.4; Fig. 1*b*).

The RoG model can capture the typical relation between response mean and variance
Cortical responses are characterized by a monotonic relation between the variance and mean of the spike count, known as Poisson-like variability (Tolhurst et al., 1983; Goris et al., 2014). Most existing models include this relation explicitly (Paninski, 2004; Pillow et al., 2008; Ecker et al., 2014; Goris et al., 2014; Lin et al., 2015; Charles et al., 2018). Differently from those models, the RoG encompasses a broader range of mean-variance relations; therefore, we first asked whether the RoG model too can capture Poisson-like variability. To constrain the RoG, we reasoned that, if N and D describe signals of neural origin, each one of them should be Poisson-like (in particular, D is usually thought of as the summed activity of many neurons, so also Poisson-like). Therefore, we assumed a power-law relation between the mean and variance (Tolhurst et al., 1983; Carandini et al., 1997) for both N and D. With these choices, the RoG model also displayed Poisson-like variability across a broad range of values for the power-law parameters, and the parameters could be easily tuned to obtain variance-to-mean relations qualitatively consistent with those typically observed in V1 (Fig. 1*c*).

Therefore, the RoG model can be easily constrained to capture Poisson-like variability and has the advantage, over other models, that it models explicitly the divisive normalization form of cortical firing rates.

Contrast modulation of V1 variability is best described by the RoG model
To test the RoG model quantitatively, we recorded responses of macaque V1 neurons ($n = 172$) to sinusoidal gratings drifting in

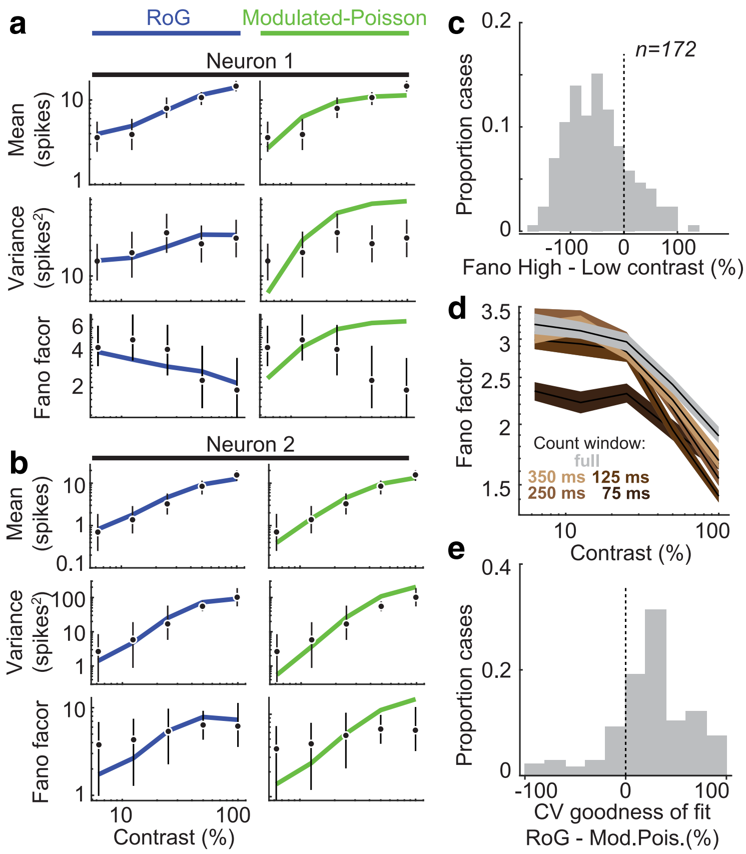

Figure 2. RoG model captures contrast modulation of firing rate and variability. ***a***, ***b***, Mean spike count (top), variance (middle), and Fano factor (bottom) as a function of contrast for two example neurons. Circles represent data. Black lines indicate 95% CI. Blue lines indicate RoG model fits. Green lines indicate modulated Poisson model fits. ***c***, Histogram of percentage difference between Fano factors at high versus low contrast. ***d***, Geometric mean (black lines) and 95% CI (shaded areas) of the Fano factor across neurons, as a function of contrast. Different colors correspond to different spike-count windows. ***e***, Percentage difference in cross-validated goodness of fit between the RoG and the modulated Poisson models. For details on model fitting and evaluation, see Materials and Methods.

4 directions at 5 contrast levels between 6.25% and 100%. Each grating was presented for 500–800 ms and repeated 20–25 times. As our interest is on the relationship between normalization and variability, for each neuron we analyzed only the responses to the best direction out of those presented, and we focused instead on contrast manipulation because normalization is known to provide excellent fits to the contrast response function (Heeger, 1992; Tolhurst and Heeger, 1997). We therefore parametrized the means of N and D in our model as in the standard normalization model (Eq. 1.16). Namely, we assumed N proportional to the squared contrast, and D equal to the squared contrast plus a constant. We then optimized jointly the parameters for the means of N and D, as well as those for the power-law relation between the means and variances, by maximum likelihood. Figure 2*a*, *b* shows for two example neurons that the RoG model captured accurately both the mean count (as expected) as well as the dependence of the variance on contrast. We quantified the predictive power of the model by the cross-validated goodness of fit (details in Materials and Methods), that is, the log-likelihood of a test dataset not used for model training, normalized between a null model (goodness of fit = 0) and the oracle model (goodness of fit = 1). Across all neurons, the median cross-validated goodness of fit was 0.85 (CI [0.80 0.87]), denoting excellent predictive power.

We observed a marked decrease in Fano factor at higher contrasts in the example neuron of Figure 2*a*, also consistent with a previous report (Orbán et al., 2016). This was because the variance increased with contrast more slowly than the mean rate (Fig. 2*a*). Across neurons, high-contrast stimuli elicited lower Fano factors compared with low-contrast stimuli for 138 of 172 neurons (average Fano at low contrast 3.2, CI [2.9 3.5]; at high contrast 1.9, CI [1.7 2.0]), and the average Fano decreased monotonically with contrast (Fig. 2*c*,*d*). We verified that the contrast dependence of the Fano factor was qualitatively similar for shorter count windows, up to 75 ms, although the average Fano was generally smaller for shorter windows (Fig. 2*d*). This behavior was captured by the RoG model (Fig. 2*a*, blue curves), although we note that the model could also capture the opposite effect in the small subset of neurons ($n = 34$ of 172) that displayed it (e.g., Fig. 2*b*). The RoG can capture both increases and decreases of Fano factor with contrast because contrast affects the means of both N and D, which can have opposing effects on Fano (Eq. 1.8). In addition, the subset of neurons whose Fano factor increased with contrast had lower spontaneous Fano factors and maximum firing rate, and larger best-fit exponents for the power law relations of N and D (the other parameters did not change substantially; Table 1).

The main effect of contrast on Fano factors was qualitatively at odds with the Poisson model (constant Fano of 1), and also with a popular extension, the modulated Poisson model (Goris et al., 2014). In the modulated Poisson model, the variance increases faster than the mean; and therefore, the Fano factor could only increase with contrast, contrary to what we observed in the data (Fig. 2*c*,*d*). To quantify this discrepancy, we fitted an alternative model in which the mean was parametrized as in the RoG model, but the variability was modulated Poisson (see Alternative models). As expected, the alternative model could not describe well the variance data, particularly when the Fano factor decreased with contrast (e.g., Fig. 2*a*, green curves), and the fit quality (median cross-validated goodness of fit 0.62, CI [0.59 0.63]) was worse than the RoG model for most neurons (Fig. 2*e*). The RoG achieved a higher fit quality (0.79 vs 0.66, $t_{(33)} = 2.39$, $p = 0.023$, paired t test) also for the small subset of neurons whose Fano factor increased with contrast (e.g., Fig. 2*b*). We note that the modulated Poisson model may be easily rescued with additional free parameters (i.e., assuming that the variance of the gain modulator is itself contrast-dependent and is reduced by high-contrast stimulation) (Goris et al., 2017).

These results indicate that the RoG model successfully extends divisive normalization, to capture simultaneously the stimulus dependence of spike count mean and variance. Therefore, this model allows us to study directly how normalization affects variability.

Table 1. Comparison of neurons with Fano factor increasing versus decreasing with contrast[a]

	Full-contrast Fano	Spontaneous Fano	R_{max}	ϵ	α_N	α_N	β_N	β_D
Fano increasing with contrast ($n = 34$)	2.40	1.55	5.8	54.8	11.3	12.6	1.70	1.60
Fano decreasing with contrast ($n = 138$)	1.87	3.56	10.9	38.7	8.7	10.1	1.53	1.30
p value (different means)	0.04	2×10^{-10}	0.0003	0.02	0.13	0.11	0.003	9×10^{-6}

[a] p values are for the difference between the means of the two subpopulations. For Fano factors, p values are computed on the logarithm. Because these analyses use the best-fit parameters of the RoG model, they were performed on the subsets of neurons with goodness of fit >0.5 (32 of 34 neurons for the first row; 120 of 138 neurons for the second row).

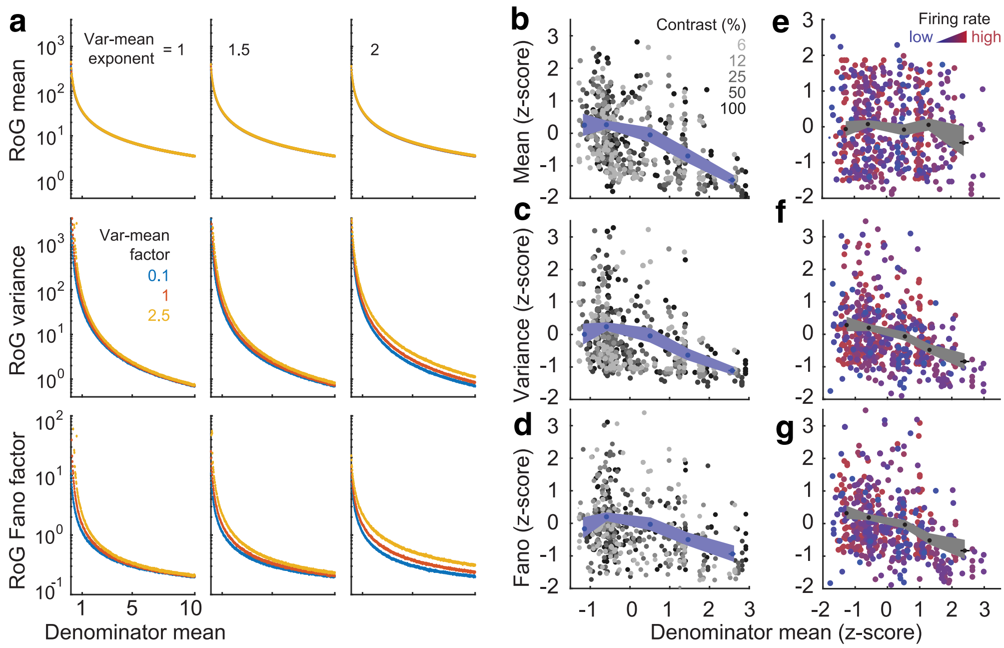

Figure 3. Normalization reduces response variability. ***a***, Relation between the normalization strength (abscissa) and mean (ordinate, top), variance (middle), or Fano factor (bottom) in the RoG model. Columns represent different values of the exponent of the relation between mean and variance of the denominator. Different colors represent different values of the proportionality factor between mean and variance of the denominator. Other parameters are set to $\mu_N = 35$; $\mu_D \in [0.1, 10]$; $\alpha_N = 1$; $\beta_N = 1$; $\rho = 0$; $\mu_\eta = 0$; $\sigma_\eta^2 = 0.1\mu_N/\mu_D$. ***b–d***, Relation between the *z*-scored normalization strength and the *z*-scored response mean (***b***), variance (***c***), or Fano factor (***d***) in the data. *z* scoring was performed across all neurons, but separately within each contrast condition. Gray symbols represent individual neurons and contrasts. Blue symbols represent average across neurons and contrasts. Shaded areas represent 95% CI. ***e–g***, Similar to ***b–d***, but with *z* scoring performed across all neurons and contrast levels with similar firing rate. Firing rate bins were logarithmically spaced between the minimum and maximum firing rates measured. Black symbols represent average across neurons and contrasts. Shaded areas represent 95% CI

Normalization often reduces response variability

We have shown above that increasing stimulus contrast often decreases Fano factors. Moreover, in the standard normalization formulation, increasing stimulus contrast also increases the strength of the normalization signal. We then asked whether this inverse relation between strength of normalization and Fano factor is general (i.e., whether increasing normalization always contributes to stabilizing the output of the RoG model). The explicit normalization form of the RoG model allowed us to answer this question by studying directly the relation between normalization and variability.

We first analyzed, in the model, the dependence of the RoG variance on the mean of D, when everything else is held constant (including the variance of D). We found analytically that, if N and D are uncorrelated, the RoG variance can only decrease when the mean of D increases (Eq. 1.4). To understand this, consider first a constant D: in this case, D simply scales the Gaussian variable N; and therefore, it scales its variance. For variable D, the RoG variance depends also on the variance of D, but the effect of the latter is small as we have assumed that D has small variance. Increasing the mean of D also decreases the RoG mean; therefore, in principle, an increase in D could have different effects on the Fano factor. However, we found analytically that the RoG mean always decreased slower than the variance, and so increasing the mean of D always reduced the Fano factor (Eq. 1.6). We further verified that these results hold also if N and D are correlated, except for extremely large correlation values (Eq. 1.7). We thus conclude that increasing normalization strength, when everything else is constant, usually stabilizes the RoG output.

We then considered the case with power-law relation between the mean and variance for N and D. In this case, increasing the mean of D increases its variance, and so the effects on the RoG variance and Fano need not be the same as in the previous analysis. Because no simple analytical solutions could be found for this case, we ran simulations in which we varied the mean of D systematically while keeping the mean of N constant, and tested different values of the parameters of the variance-to-mean relation for N and D. We observed a strong trend for reduced Fano factors as the mean of D increased, across a broad range of parameter settings (Fig. 3*a*), although this trend could be reversed for larger values of the exponent in the power law of D, or if the additive noise has a relatively large Fano factor (see Eqs. 1.8 and 1.9 and related text).

This analysis thus suggests that, in the RoG model, normalization often reduces spike count variability and Fano factors. If this is the case also in V1, we would expect that, in our dataset, neurons that are more strongly normalized are also less variable. To test the prediction, we first computed for each neuron and each

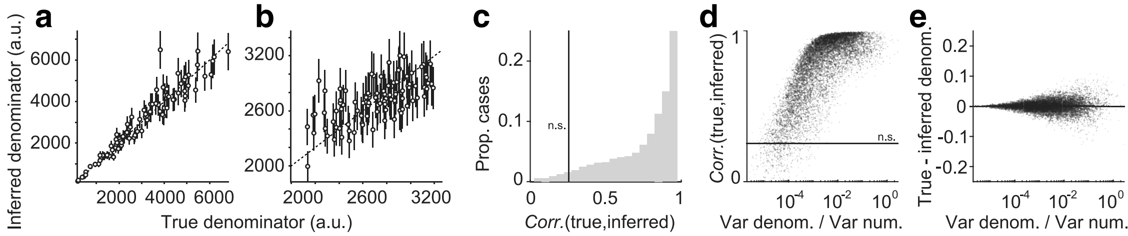

Figure 4. Inference of single-trial normalization strength in the RoG model. ***a***, ***b***, Comparison of the true and inferred single-trial normalization strength, across 100 trials of two simulated experiments. Error bars: standard deviation of the inferred values. ***c***, Histogram of Spearman correlation coefficients between true and inferred values, across 10,000 simulated experiments each with 100 trials. The model was formulated as in Methods eq. (1.16) for the contrast response function, with parameters drawn from the uniform distribution in the following intervals: $R_{\max} \in [10,100]$; $\epsilon \in [15,25]$; $\beta_{\mathrm{N}} = \beta_{\mathrm{D}} \in [1.5,2]$; $\rho = 0$; $\mu_{\eta} = 0$; $\sigma_{\eta}^{2} = 0$, with contrast between 20 and 50%, and with α_N, α_D set to enforce a Fano factor of 1 at a 75% contrast. ***d***, Correlation between true and inferred values (ordinate) as a function of the ratio between the variances of the denominator and numerator in the RoG model (abscissa). Correlation coefficients smaller than the vertical line (***c***) or horizontal line (***d***) are not statistically significant ($p > 0.05$). ***e***, Difference between true and inferred values of the denominator, expressed as a fraction of the true value.

contrast level the mean of D (i.e., the normalization strength) from the best fit parameters. Because the analysis relies on the fit parameters, we only included the 152 of 172 neurons with goodness of fit > 0.5. Different from the model analysis above, in which we varied D while keeping N constant, in the V1 data, both D and N varied across neurons and contrast levels. Therefore, to isolate the effect of D, we performed two complementary analyses. First, we analyzed each contrast level separately and we z-scored the strength of D across neurons. We found that neurons with stronger normalization tended to have lower mean as expected (Fig. 3*b*) and even lower variance (Fig. 3*c*), resulting often in lower Fano factors (Fig. 3*d*). Across neurons and contrast levels, there was a significant negative correlation between the z score of Fano and of D (Spearman correlation $= -0.29$, $p = 6 \times 10^{-16}$). The negative correlation was significant also at each individual contrast level ($p < 0.03$), except at 50% contrast ($p = 0.68$). Second, we used mean-matching (Churchland et al., 2010) to assess the relation between normalization strength and variability, independent of firing rate. Specifically, we binned, across neurons and contrast levels, all cases with similar firing rate and then computed z scores within each bin. In this case, as expected, there was no relation between the firing rate and D (Fig. 3*e*; Spearman correlation $= -0.02$, $p = 0.6$), but we observed a strong negative correlation between the variance and D (Fig. 3*f*; Spearman correlation $= -0.32$, $p = 3 \times 10^{-19}$) and between the Fano and D (Fig. 3*g*; Spearman correlation $= -0.38$, $p = 9 \times 10^{-27}$). We verified that these results remained qualitatively similar for the shorter spike-count windows of Figure 2*d*, although fewer neurons could be analyzed (see Materials and Methods). Therefore, the RoG model predicts a general, inverse relation between the average normalization strength and the variability of cortical responses, which is apparent in V1 activity.

Inference of single-trial normalization strength

The RoG model extends the classical normalization model, allowing it to simultaneously capture the mean and variance of the spike count across trials. An important advantage of the RoG formulation is that, because it treats the normalization signal explicitly as a random variable, it allows for inference of the normalization strength in single-trial activity, a quantity that cannot be measured directly with current experimental techniques. Therefore, we derived analytical expressions for the inferred value of the single-trial normalization strength (i.e., the most probable value of D for the observed spike count, according to the RoG model) and its uncertainty. Intuitively, if we knew the single-trial value of the numerator N, then we could estimate D simply as the ratio between N and the measured spike count. However, because N is not measurable either, we have to compute a so-called posterior probability distribution of the possible values of D, which amounts approximately to combining different estimates of D corresponding to all possible values of N (termed marginalization). This posterior distribution is computed via a direct application of Bayes rule, that is, combining the evidence (the likelihood of the observed spike count for different possible values of D, under the RoG model) with the prior (i.e., the probability of the expected values of D, before observing the spike count; Eq. 1.10). The prior distribution of D is, by the definition of the RoG model, a Gaussian per Equation 1.2 whose (prior) mean and variance are estimated separately by model fitting across trials as shown above. The posterior distribution can be approximated by a Gaussian too, which allowed us to derive simple closed-form expressions for its (posterior) mean and variance (Methods eq. (1.12)), corresponding respectively the inferred value of D and the associated uncertainty.

We first validated these estimators in 10,000 simulated experiments, each with a different set of parameters for N and D. The example in Figure 4*a* shows one experiment in which, across 100 simulated trials, estimates of D were strongly correlated with the ground-truth value. Across these simulated experiments, however, we observed a broad range of correlations (Fig. 4*b*,*c*), including cases with nonsignificant correlation. We reasoned that, if the prior variance of D is much smaller than that of N, the inferred single-trial value of D cannot be accurate. Intuitively, this is because, when the prior variance of D is small, the single-trial value of D is always very close to the mean (i.e., it is almost constant across trials); therefore, across-trial changes in the ratio are largely determined by across-trial changes in N. Therefore, while the average of D can be estimated accurately, its single-trial variations around the average cannot. This is also evident from Equation 1.12, where, in the limit that the variance of N is large, the single-trial estimate of D becomes equal to its prior mean and independent of the measured spike count in that trial. Conversely, when the variance of D is larger relative to N, then the inference of D should be more accurate. In agreement with this, we found in our simulations that the quality of the inferred values of D increased systematically with the ratio between the variances of D and N (Fig. 4*d*). We note, however, that the estimates remained unbiased on average (within 0.05% across all simulated experiments) regardless of the ratio of variances (Fig. 4*e*).

This analysis suggests that, for neurons and stimulus conditions with a relatively large variance of D, we can infer accurately

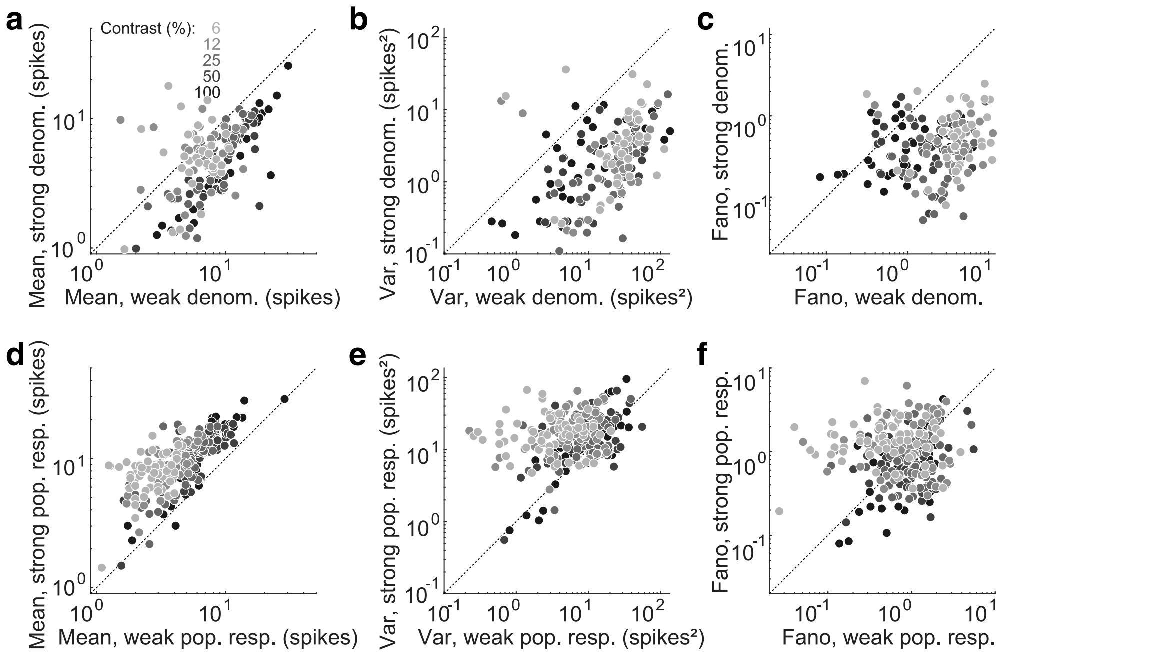

Figure 5. Reduced response variability during epochs with strong normalization. ***a***, Each symbol denotes, for one neuron and one contrast condition, the mean spike count across trials with inferred strong (ordinate) versus trials with inferred weak (abscissa) normalization signal. Only neurons with large across-trials variance of the normalization signal (at least 1% of the variance of the numerator) and trials with at least one spike are included. ***b***, ***C***, Same as ***a***, but for response variance (***b***) and Fano factor (***c***). ***d***, ***e***, ***f***, Same as ***a***, ***b***, ***c***, but using as a proxy of single-trial normalization strength the total activity of a population of simultaneously recorded neurons.

the single-trial value of D. Out of the 153/172 neurons for which the RoG fit quality was larger than 0.5, we found that $n = 151$ neurons satisfied this criterion (specifically, that the variance of D was at least 1% of the variance of N) for at least one stimulus contrast level. We also excluded from this analysis the trials with zero spikes, for which D cannot be estimated accurately (see Methods), leaving us with $n = 112$ neurons. For those cases, we predicted that, if the inference is accurate, during trials with larger normalization signal, the response variability should be lower. To test this prediction, we estimated D in each trial, and then split trials within each contrast condition in two sets containing values of D either above or below average. As a control, we verified that in the trial set with large inferred D, the mean spike count was always lower in most cases (Fig. 5*a*), indicating that our estimates of D were meaningful. Importantly, both the spike-count variance and the Fano factor were also systematically lower for the large-D trials (Fig. 5*b*,*c*), in agreement with the prediction of the RoG model. Across neurons and conditions, we observed a 142% reduction (CI [139 151]%) in Fano factor for large-D versus small-D trials. We verified that these results remained qualitatively similar for the shorter spike-count windows of Figure 2*d*, although fewer neurons could be analyzed (see Materials and Methods). This result was also robust to the choice of threshold (i.e., the minimum ratio between the variances of D and N) (Fig. 6*a*–*c*). The changes were absent if we randomly split trials in two groups (average difference in Fano >1000 splits = −5%, CI [−13, 6]%). As a further control, we verified that, if we instead considered only cases in which the ratio between the variances of D and N was less than a small threshold (i.e., where we expect poor single-trial estimates of D), there were no significant differences in Fano factor between large-D and small-D trials, unless we made the threshold large enough to effectively include many cases with relatively large variance of D (Fig. 6*d*–*f*).

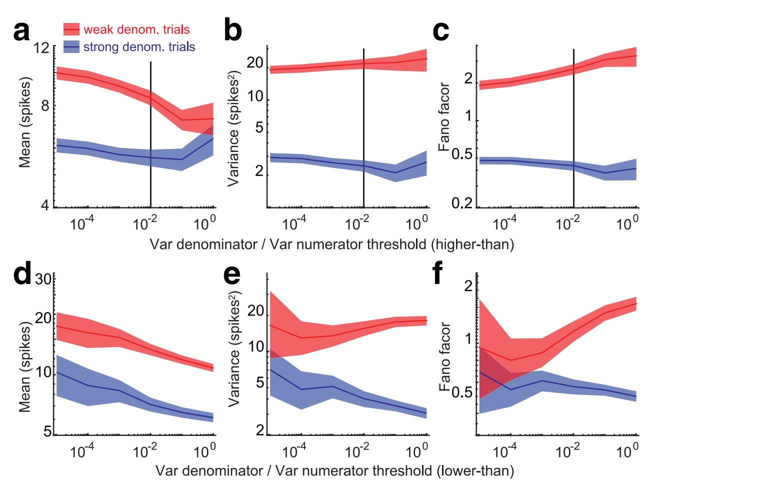

Figure 6. Reduced variability during epochs with strong inferred normalization is robust to choice of threshold. ***a***, Average spike count (ordinate) across neurons and contrast conditions for which the ratio between denominator variance and numerator variance was larger than the threshold (abscissa). Averages are computed over the subset of trials with strong (blue) or weak (red) inferred normalization signal. Shaded areas: 95% CI Black vertical line: the threshold used for Fig. 5. ***b***, ***c***, Same as (***a***) but for the spike count variance (***b***) and Fano factor (***c***). ***d–f***, Same as ***a–c***, but including neurons and contrast conditions for which the ratio between denominator variance and numerator variance was smaller than the threshold (abscissa). The largest threshold in ***d–f*** and the smallest threshold in ***a–c*** correspond to including all neurons and contrast conditions.

Because normalization is thought to result from the pooled activity of many neurons (Heeger, 1992; Carandini and Heeger,

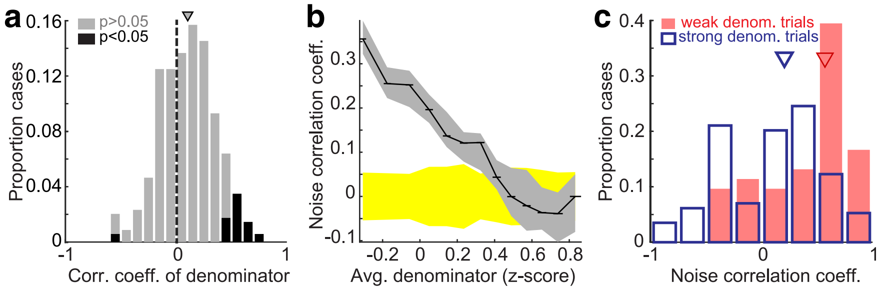

Figure 7. Normalization is shared between neurons and reduces noise correlation. ***a***, Histogram of correlation coefficients of the inferred single-trial denominator for pairs of simultaneously recorded neurons with similar tuning. Gray bars: pairs with non-significant correlation; black bars: pairs with significant correlation. Triangle: average correlation across all pairs. ***b***, Black symbols and gray shaded area: Median noise correlation and 95% CI, as a function of the average normalization strength of each pair. For each pair, the inferred values of *D* of each neuron were z-scored across trials; then trials were sorted by the average z-scored *D* of the pair. Each point in the plot represents averages computed in windows of 12 trials. Yellow area: 95% CI for a control with trial order shuffled independently for each neuron, to remove noise correlations. ***c***, Histogram of noise correlation coefficients during trials with large (blue) or small (red) inferred normalization signal. These estimates of correlations are based on few trials (as little as 5 in some cases, see related text), hence the large range of values. Triangles denote group medians.

2011), we reasoned that the normalization signal should covary between neurons on a trial-by-trial basis. To test this hypothesis, we measured the correlation between the inferred denominator for pairs of neurons recorded simultaneously. For each neuron, we only analyzed responses to the grating orientation that elicited maximal activity (see Materials and Methods); therefore, for this analysis, we considered pairs of neurons with similar tuning preference ($n = 179$ pairs). We found that D was on average positively correlated (median correlation 0.1 across 344 pairs and contrast levels; Fig. 7*a*), and that most cases with a significant correlation were positive (median correlation 0.54 across 27 pairs and contrast levels with $p < 0.05$; Fig. 7*a* black bars). Because it reflects a shared signal, normalization is also thought to affect the coordination between neurons (Verhoef and Maunsell, 2017). Our model allowed us to test this directly, by relating pairwise noise correlations to normalization strength. First, for each pair of neurons, we sorted the trials by the average denominator (i.e., the average between the two neurons of the z-scored values of D) and found that noise correlations decreased approximately monotonically as the average denominator increased (Fig. 7*b*). Next, for each pair and contrast, we separated the trials with large D for both neurons and those with small D for both neurons ($n = 114$ cases had at least 5 trials of each kind); noise correlations were substantially reduced across trials with large denominator relative to those with weak denominator (Fig. 7*c*).

We next asked whether the total population activity may be a good proxy for the single-trial normalization strength, under the assumption that normalization reflects the pooled activity of many neurons. However, when we split trials based on total population activity, we found larger mean spike count for trials with larger population activity (an average increase of 70%, CI [65 75]%; opposite to what would be predicted by stronger normalization) but no significant change in Fano factor (a 7% increase, CI [−6 21]%; Fig. 5*d–f*). This may reflect that, whereas the normalization strength is conceptualized as the total activity of the un-normalized population, the measured population activity is already normalized, and it is therefore not a good measure of normalization strength. Alternatively, it may be that the normalization pool is distinct, or much larger, than the neurons we could record from simultaneously.

These results thus demonstrate that the RoG model enables accurate inference of the single-trial normalization strength for the majority of neurons. Using this method, we found that response variability, as well as noise correlations between similarly tuned neurons, were lowered during epochs with strong normalization.

Discussion

Divisive normalization is ubiquitous in neural computation and can dramatically affect both the strength and the variability of neuronal activity. Existing models do not offer a way to quantify the influence of normalization on variability. Here we have introduced a simple yet powerful model, the RoG, that extends the standard normalization model to variable neuronal activity. The RoG offers two major advantages over existing models of variability: (1) the explicit normalization formulation allowed us to study directly the impact of normalization on response variability; and (2) the RoG treats the normalization signal as a hidden random variable, which allowed us to estimate the single-trial normalization strength via Bayesian inference.

The logic of the model proposed here is fundamentally different from that of models proposed by us and others to explain the widespread presence of normalization (Schwartz and Simoncelli, 2001; Wainwright et al., 2001; Schwartz et al., 2007; Beck et al., 2011; Coen-Cagli et al., 2012, 2015; Lochmann et al., 2012; Sinz and Bethge, 2013; Snow et al., 2016). Models in this class hypothesize a specific goal of neurons (e.g., efficient coding of sensory inputs) (Barlow, 1961) and ask whether normalization achieves that goal (i.e., these models aim to explain why neurons behave in a way that is well described by normalization). Our goal here instead was to provide a descriptive model (i.e., a model that parametrizes the effects of normalization on neuronal responses in a way that is easy to fit to data), allowing us to quantify and interpret those effects. Using this approach, we have quantified the relation between normalization and variability; explaining why this relation exists and what functional role it may play, falls in the domain of the other class of models (see, e.g., Orbán et al., 2016).

The RoG described accurately the contrast dependence of spike count mean and variance in macaque V1. In particular, the model was able to reproduce the systematic reduction of the Fano factor as contrast increased, which we observed in the majority of recorded neurons (Fig. 2*c*,*d*). This effect is consistent with similar recent observations (Orbán et al., 2016), but it could not be captured by the popular modulated Poisson model of response variability (Goris et al., 2014), although this model may be rescued with additional free parameters (Goris et al., 2017) (Fig. 2*e*). Another recent model termed flexible overdispersion (Charles et al., 2018) extends the classical Linear-Nonlinear-Poisson (LNP) framework, and can capture a decrease of Fano factor as firing rate increases, and thus could fit well our observation on contrast response functions. However, this model uses a simple response nonlinearity as is common in the LNP framework for tractable optimization, and so it would not allow to study the role of divisive normalization. One possibility would be to treat both the numerator and denominator of the standard normalization model as LNPs (i.e., similar to the RoG but using LNPs instead of Gaussians). The disadvantage of this approach is that the spike-count distribution may not be well approximated by a Gaussian, and one would no longer have simple expressions for the spike count mean and variance as in Equation 1.4, nor closed-form, instantaneous inference of the single-trial normalization signal as in Equation 1.12. A related point is that, when very few spikes are observed (e.g., using short count windows, see Materials and Methods), the RoG does not fit the data well because the spike-count distribution deviates markedly from Gaussian. This could be mitigated using a Poisson distribution for the spike count conditional on the ratio *D/N*, and gamma distributions on *D* and *N*. However, the resulting distribution (Holla and Bhattacharya, 1965; Karlis and Xekalaki, 2005), termed mixed Poisson beta distribution of the second kind, has the same limitations mentioned above, and furthermore cannot capture Fano factors <1.

Thanks to the explicit normalization in the RoG, analysis of the model indicated a general link between normalization and variability (Fig. 3*a*). Consistent with this, we observed that neurons with stronger normalization were relatively less variable (Fig. 3*b*–*d*), and also that, for most neurons, variability was reduced during response epochs in which normalization was stronger than average (Fig. 5*a*–*c*). One possible interpretation of these observations is that normalization may reflect a general mechanism to stabilize cortical activity, for instance, one based on dynamic balance between excitation and inhibition (Adesnik, 2017; Hennequin et al., 2018). However, diverse influences of normalization on variability may be apparent across cortical areas or with different experimental manipulations, possibly reflecting the diverse network mechanisms thought to underlie the effects described by normalization (Carandini and Heeger, 2011), including feedforward (Carandini et al., 2002; Webb et al., 2005), recurrent (Ozeki et al., 2009; Rubin et al., 2015), and feedback (Nassi et al., 2013; Nurminen et al., 2018) signals. Although our analysis revealed primarily a stabilizing effect of normalization, the RoG model is flexible enough to capture also the opposite effect.

Two recent studies have also pointed to a possible impact of normalization on response variability (Tripp, 2012; Verhoef and Maunsell, 2017). Those studies focused on variability shared between neurons (i.e., noise correlations) and relied exclusively on simulations. In contrast, the RoG framework allowed us to derive analytical results linking normalization and variability. Furthermore, the RoG allows for quantitative data fitting and single-trial inference, which are not tractable in the models of others (Tripp, 2012; Verhoef and Maunsell, 2017). The RoG model could also be extended to capture the joint responses of two or more neurons, simply by including *N* and *D* terms for each neuron, and assuming a multivariate Gaussian distribution for the resulting collection of variables (i.e., with dimension 2 times the number of neurons). With this extension, the number of parameters (i.e., the covariance matrix of the multivariate Gaussian) would grow quadratically with the number of neurons, thus potentially limiting the applicability or requiring strong regularization for large numbers of neurons. However, for pairwise responses and under the assumption of independence between *N* and *D* (which was supported by our data, see Materials and Methods), the pairwise RoG would only require two additional parameters for the correlation between the *N* terms and between the *D* terms, plus one for spontaneous noise correlations.

To optimize the RoG parameters on data, it is necessary to consider stimulus manipulations for which the firing rate is well described by the standard normalization model. This is the case for contrast response functions, which motivated our choice here to validate the RoG model. There are other stimulus manipulations that are thought to involve normalization in V1, as observed in phenomena, such as cross-orientation masking (Bonds, 1989; DeAngelis et al., 1992; Carandini et al., 1997; Schwartz and Simoncelli, 2001), surround suppression (Schwartz and Simoncelli, 2001; Cavanaugh et al., 2002; Coen-Cagli et al., 2015; Trott and Born, 2015), and temporal adaptation (Heeger, 1992; Solomon and Kohn, 2014; Snow et al., 2016; Westrick et al., 2016). Furthermore, normalization is also widespread (Carandini and Heeger, 2011) in other visual areas (Rust et al., 2006), and in other sensory (Olsen et al., 2010; Ohshiro et al., 2011; Rabinowitz et al., 2011) and nonsensory (Louie et al., 2014) areas. Applying the RoG model to those datasets could therefore reveal whether normalization has an impact on variability in other domains, similar to what we found in V1 for contrast.

One case of particular interest is temporal adaptation, as it has been hypothesized that adaptation effects reflect a modulation of the normalization signal itself (Heeger, 1992; Wainwright et al., 2001; Solomon and Kohn, 2014; Westrick et al., 2016; Aschner et al., 2018). Inference of the single-trial normalization signal, as enabled by the RoG model, would allow testing this hypothesis directly. However, one limitation of the single-trial inference of normalization strength is that it is only reliable on trials with at least one spike, and for neurons and contrast levels with a relatively large across-trial variance of the normalization signal. In our dataset, the normalization variance was larger than 1% of the numerator variance in $<70\%$ of the cases (516 of 760 neurons and contrast levels). This percentage may reflect that the effective normalization signal results from the pooled activity of many neurons, as commonly assumed in the normalization framework (Carandini and Heeger, 2011) and further supported by our finding that the single-trial normalization signal is positively correlated between neurons (Fig. 7*a*). If neurons in the normalization pool have homogeneous response properties, pooling their responses should produce a more stable signal than the excitatory drive to one neuron. This explanation also suggests that neurons with a relatively larger variance of *D*, according to the RoG, may be normalized by a more heterogeneous normalization pool. This could be tested by measuring the tuning of the normalization signals (Webb et al., 2005) and studying whether normalization tuning width correlates with the estimated variance of the normalization signal. Furthermore, it should be possible to differentially adapt the normalization pools of the two classes of neurons: for instance, a narrow-band adapter would weaken normalization effects more for the neurons with small variance of *D* according to the RoG.

Single-trial inference of normalization could also prove useful in understanding attentional modulation of perception. It has been suggested that attention improves perceptual decision-making by modulating correlated response variability (Cohen and Maunsell, 2009; Mitchell et al., 2009), and that this modulation results from a dynamic mechanism for normalization (Verhoef and Maunsell, 2017). Consistent with the results of Verhoef and Maunsell (2017), our comparison of single-trial inferences across neurons revealed that noise correlations between pairs of similarly tuned neurons are reduced during epochs of strong normalization (Fig. 7*b*,*c*). Our model (or an extension that explicitly captures pairwise responses and noise correlations) could allow one to test this hypothesis directly by relating single-trial normalization strength and behavioral performance.

References

Adesnik H (2017) Synaptic mechanisms of feature coding in the visual cortex of awake mice. Neuron 95:1147–1159.e4.

Albrecht DG, Geisler WS (1991) Motion selectivity and the contrast-response function of simple cells in the visual cortex. Vis Neurosci 7:531–546.

Arieli A, Sterkin A, Grinvald A, Aertsen A (1996) Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273:1868 –1871.

Aschner A, Solomon SG, Landy MS, Heeger DJ, Kohn A (2018) Temporal contingencies determine whether adaptation strengthens or weakens normalization. J Neurosci 38:10129 –10142.

Barlow HB (1961) Possible principles underlying the transformation of sensory messages. In: Sensory communication (Rosenblith WA, ed), pp 217–234. Cambridge, MA: Massachusetts Institute of Technology.

Beck JM, Latham PE, Pouget A (2011) Marginalization in neural circuits with divisive normalization. J Neurosci 31:15310 –15319.

Bennett C, Arroyo S, Hestrin S (2013) Subthreshold mechanisms underlying state-dependent modulation of visual responses. Neuron 80:350 –357.

Bonds AB (1989) Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. Vis Neurosci 2:41–55.

Carandini M, Heeger DJ (2011) Normalization as a canonical neural computation. Nat Rev Neurosci 13:51– 62.

Carandini M, Heeger DJ, Movshon JA (1997) Linearity and normalization in simple cells of the macaque primary visual cortex. J Neurosci 17:8621–8644.

Carandini M, Heeger DJ, Senn W (2002) A synaptic explanation of suppression in visual cortex. J Neurosci 22:10053–10065.

Cavanaugh JR, Bair W, Movshon JA (2002) Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. J Neurophysiol 88:2530 –2546.

Charles AS, Park M, Weller JP, Horwitz GD, Pillow JW (2018) Dethroning the Fano factor: a flexible, model-based approach to partitioning neural variability. Neural Comput 30:1012–1045.

Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM, Moore T, Chang SW, Snyder LH, Lisberger SG, Priebe NJ, et al. (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat Neurosci 13:369 –378.

Coen-Cagli R, Dayan P, Schwartz O (2012) Cortical surround interactions and perceptual salience via natural scene statistics. PLoS Comput Biol 8:e1002405.

Coen-Cagli R, Kohn A, Schwartz O (2015) Flexible gating of contextual influences in natural vision. Nat Neurosci 18:1648 –1655.

Cohen MR, Maunsell JH (2009) Attention improves performance primarily by reducing interneuronal correlations. Nat Neurosci 12:1594 –1600.

Dadarlat MC, Stryker MP (2017) Locomotion enhances neural encoding of visual stimuli in mouse V1. J Neurosci 37:3764 –3775.

DeAngelis GC, Robson JG, Ohzawa I, Freeman RD (1992) Organization of suppression in receptive fields of neurons in cat visual cortex. J Neurophysiol 68:144 –163.

Ecker AS, Berens P, Cotton RJ, Subramaniyan M, Denfield GH, Cadwell CR, Smirnakis SM, Bethge M,Tolias AS (2014) State dependence of noise correlations in macaque primary visual cortex. Neuron 82:235–248.

Goris RL, Movshon JA, Simoncelli EP (2014) Partitioning neuronal variability. Nat Neurosci 17:858 – 865.

Goris RL, Meding K, Henaff OJ (2017) Representation of sensory uncertainty in Macaque visual cortex. Society for Neuroscience, Washington, DC.

Harris KD, Thiele A (2011) Cortical state and attention. Nat Rev Neurosci 12:509 –523.

Heeger DJ (1992) Normalization of cell responses in cat striate cortex. Vis Neurosci 9:181–197.

Hennequin G, Ahmadian Y, Rubin DB, Lengyel M, Miller KD (2018) The dynamical regime of sensory cortex: stable dynamics around a single stimulus-tuned attractor account for patterns of noise variability. Neuron 98:846 – 860.e5.

Hinkley DV (1969) On the ratio of two correlated normal random variables. Biometrika 56:635– 639.

Holla MS, Bhattacharya SK (1965) On a discrete compound distribution. Ann Inst Statist Math 15:377–384.

Jia X, Tanabe S, Kohn A (2013) Gamma and the coordination of spiking activity in early visual cortex. Neuron 77:762–774.

Karlis D, Xekalaki E (2005) Mixed Poisson distributions. Int Statist Rev 73:35–58.

Lin IC, Okun M, Carandini M, Harris KD (2015) The nature of shared cortical variability. Neuron 87:644 – 656.

Lochmann T, Ernst UA, Denève S (2012) Perceptual inference predicts contextual modulations of sensory responses. J Neurosci 32:4179 – 4195.

Louie K, LoFaro T, Webb R, Glimcher PW (2014) Dynamic divisive normalization predicts time-varying value coding in decision-related circuits. J Neurosci 34:16046 –16057.

Marsaglia G (1965) Ratios of normal variables and ratios of sums of uniform variables. J Am Statist Assoc 60:193–204.

Marsaglia G (2006) Ratios of normal variables. J Statist Softw 1:4.

Mitchell JF, Sundberg KA, Reynolds JH (2009) Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. Neuron 63:879 – 888.

Nassi JJ, Lomber SG, Born RT (2013) Corticocortical feedback contributes to surround suppression in V1 of the alert primate. J Neurosci 33:8504 – 8517.

Nurminen L, Merlin S, Bijanzadeh M, Federer F, Angelucci A (2018) Top-down feedback controls spatial summation and response amplitude in primate visual cortex. Nat Commun 9:2281.

Ohshiro T, Angelaki DE, DeAngelis GC (2011) A normalization model of multisensory integration. Nat Neurosci 14:775–782.

Olsen SR, Bhandawat V, Wilson RI (2010) Divisive normalization in olfactory population codes. Neuron 66:287–299.

Orbán G, Berkes P, Fiser J, Lengyel M (2016) Neural variability and sampling-based probabilistic representations in the visual cortex. Neuron 92:530 –543.

Ozeki H, Finn IM, Schaffer ES, Miller KD, Ferster D (2009) Inhibitory stabilization of the cortical network underlies visual surround suppression. Neuron 62:578 –592.

Paninski L (2004) Maximum likelihood estimation of cascade point-process neural encoding models. Network 15:243–262.

Pham-Gia T, Turkkan N (2005) Distributions of ratios of random variables from the power-quadratic exponential family and applications. Statistics 39:355–372.

Pham-Gia T, Turkkan N, Marchand E (2006) Density of the ratio of two normal random variables and applications. Commun Statist Theory Methods 35:1569 –1591.

Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, Simoncelli EP (2008) Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature 454:995–999.

Ponce-Alvarez A, Thiele A, Albright TD, Stoner GR, Deco G (2013) Stimulus-dependent variability and noise correlations in cortical MT neurons. Proc Natl Acad Sci U S A 110:13162–13167.

Rabinowitz NC, Willmore BD, Schnupp JW, King AJ (2011) Contrast gain control in auditory cortex. Neuron 70:1178 –1191.

Rabinowitz NC, Goris RL, Cohen M, Simoncelli EP (2015) Attention stabilizes the shared gain of V4 populations. eLife 4:e08998.

Reynolds JH, Heeger DJ (2009) The normalization model of attention. Neuron 61:168 –185.

Rosen MI (1995) Niels Hendrik Abel and equations of the fifth degree. Am Math Monthly 102:495–505.

Rubin DB, Van Hooser SD, Miller KD (2015) The stabilized supralinear network: a unifying circuit motif underlying multi-input integration in sensory cortex. Neuron 85:402– 417.

Rust NC, Mante V, Simoncelli EP, Movshon JA (2006) How MT cells analyze the motion of visual patterns. Nat Neurosci 9:1421–1431.

Schwartz O, Simoncelli EP (2001) Natural signal statistics and sensory gain control. Nat Neurosci 4:819 – 825.

Schwartz O, Hsu A, Dayan P (2007) Space and time in visual context. Nat Rev Neurosci 8:522–535.

Sinz F, Bethge M (2013) Temporal adaptation enhances efficient contrast gain control on natural images. PLoS Comput Biol 9:e1002889.

Snow M, Coen-Cagli R, Schwartz O (2016) Specificity and timescales of

cortical adaptation as inferences about natural movie statistics. J Vis 16:1.
Snyder AC, Morais MJ, Kohn A, Smith MA (2014) Correlations in V1 are reduced by stimulation outside the receptive field. J Neurosci 34: 11222–11227.
Solomon SG, Kohn A (2014) Moving sensory adaptation beyond suppressive effects in single neurons. Curr Biol 24:R1012–R1022.
Springer MD (1979) The algebra of random variables. New York: Wiley.
Stocker AA, Simoncelli EP (2006) Noise characteristics and prior expectations in human visual speed perception. Nat Neurosci 9:578–585.
Tolhurst DJ, Heeger DJ (1997) Comparison of contrast-normalization and threshold models of the responses of simple cells in cat striate cortex. Vis Neurosci 14:293–309.
Tolhurst DJ, Movshon JA, Dean AF (1983) The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision Res 23:775–785.
Tripp BP (2012) Decorrelation of spiking variability and improved information transfer through feedforward divisive normalization. Neural Comput 24:867–894.
Trott AR, Born RT (2015) Input-gain control produces feature-specific surround suppression. J Neurosci 35:4973–4982.
Verhoef BE, Maunsell JH (2017) Attention-related changes in correlated neuronal activity arise from normalization mechanisms. Nat Neurosci 20:969–977.
Wainwright MJ, Schwartz O, Simoncelli EP (2001) Natural image statistics and divisive normalization: modeling nonlinearity and adaptation in cortical neurons. In: Probabilistic models of the brain: perception and neural function, pp 203–222. Cambridge, MA: Massachusetts Institute of Technology.
Webb BS, Dhruv NT, Solomon SG, Tailby C, Lennie P (2005) Early and late mechanisms of surround suppression in striate cortex of macaque. J Neurosci 25:11666–11675.
Westrick ZM, Heeger DJ, Landy MS (2016) Pattern adaptation and normalization reweighting. J Neurosci 36:9805–9816.